PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

D6

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00 | A2

(11) International Publication Number: WO 98/42741

(43) International Publication Date: 1 October 1998 (01.10.98)

(21) International Application Number: PCT/US98/05972

(22) International Filing Date: 25 March 1998 (25.03.98)

(30) Priority Data:

08/825,145	25 March 1997 (25.03.97)	US
09/046,881	24 March 1998 (24.03.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Street, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

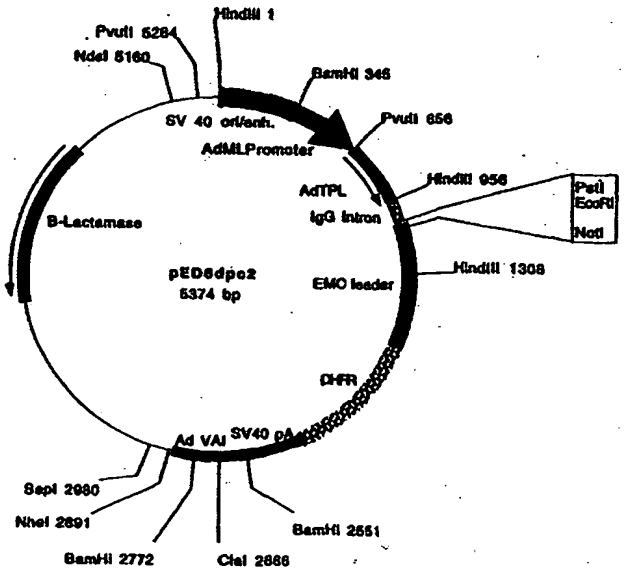
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

RAPPORT DE RECHERCHE INTERNATIONALE

Dem. : Internationale No
PCT/FR 96/02061

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	DATABASE EMBL ID: HS152227, AC= H72152, 2 Novembre 1995 XP002019920 voir l'alignements des séquences nucléotidiques et protéiques; les descripteurs & UNPUBLISHED, 1995, HILLIER ET AL.: ---	1,3,7, 17,26
A,P	DATABASE EMBL ID: HS49264, AC=N31049, 12 Janvier 1996 XP002019921 Voir l'alignement des séquences nucléotidiques et protéiques; les descripteurs. & UNPUBLISHED, 1996, HILLIER ET AL.: ---	1,3,7, 17,26
P,X	PROC. NATL ACAD. SCI., vol. 93, no. 9, 30 Avril 1996, pages 3953-3957, XP002032914 AMSON ET AL.: "Isolation of 10 differentially expressed cDNAs in p53-induced apoptosis : activation of the vertebrate homologue of Drosophila seven in absentia gene" voir le document en entier ---	1-10,17, 19,21, 24,26,27
P,X	PROC. NATL ACAD. SCI., vol. 93, no. 17, 20 Août 1996, pages 9039-9042, XP000611649 NEMANI ET AL.: "Activation of the human homologue of the Drosophila sina gene in apoptosis and tumor suppression" voir le document en entier ---	1-3,5-7, 17,19,26
A	NATURE, vol. 375, 29 Juin 1995, pages 754-760, XP002032915 SHERRINGTON ET AL.: "Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease" cité dans la demande voir le document en entier ---	1,8,10, 25,28,29
A	AUSTRALIAN AND NEW ZEALAND JOURNAL OF MEDICINE, vol. 25, no. 6, Décembre 1995, pages 845-851, XP000610669 DELLA N G ET AL: "A COMBINED GENETIC AND BIOCHEMICAL APPROACH TO MAMMALIAN SIGNAL TRANSDUCTION" voir le document en entier -----	1,3,7

1

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Dem. Internationale No
PCT/FR 96/02061

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9519367 A	20-07-95	US 5484710 A	16-01-96
WO 9511301 A	27-04-95	AU 7983294 A	08-05-95

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/825,145), filed March 25, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 54 to nucleotide 737;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 188 to nucleotide 671;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bf171_6 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bf171_6 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 109 to amino acid 118 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 54 to nucleotide 737; the nucleotide sequence of SEQ ID NO:1 from nucleotide 188 to nucleotide 671; the nucleotide sequence of the full-length protein coding

sequence of clone bf171_6 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone bf171_6 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 46 to amino acid 206.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 46 to amino acid 206;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 109 to amino acid 118 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 46 to amino acid 206.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 135 to nucleotide 1169;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 875;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 167 to amino acid 176 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 135 to nucleotide 1169; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 875; the nucleotide sequence of the full-length protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 247.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 247;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 167 to amino acid 176 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 247.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 882 to nucleotide 1106;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1050 to nucleotide 1106;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1028 to nucleotide 1395;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone co736_3 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone co736_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 32 to amino acid 41 of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 882 to nucleotide 1106; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1050 to nucleotide 1106; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1028 to nucleotide 1395; the nucleotide sequence of the full-length protein coding sequence of clone co736_3 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone co736_3 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 32 to amino acid 41 of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2283 to nucleotide 2858;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1164 to nucleotide 1433;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 91 to amino acid 100 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 2283 to nucleotide 2858; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1164 to nucleotide 1433; the nucleotide sequence of the full-length protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371; or the

nucleotide sequence of a mature protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 91 to amino acid 100 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 168 to nucleotide 683;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 318 to nucleotide 683;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 81 to amino acid 90 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 168 to nucleotide 683; the nucleotide sequence of SEQ ID NO:9 from nucleotide 318 to nucleotide 683; the nucleotide sequence of the full-length protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 53 to amino acid 172.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9 or SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 53 to amino acid 172;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 81 to amino acid 90 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 53 to amino acid 172.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 67 to nucleotide 879;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 118 to nucleotide 879;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1224 to nucleotide 2171;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:12 from nucleotide 67 to nucleotide 879; the nucleotide sequence of SEQ ID NO:12 from nucleotide 118 to nucleotide 879; the nucleotide sequence of SEQ ID NO:12 from nucleotide 1224 to nucleotide 2171; the nucleotide sequence of the full-length protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 119.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:12.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 119;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:13 or the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 119.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 2 to nucleotide 556;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 53 to nucleotide 556;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 367;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 87 to amino acid 96 of SEQ ID NO:15;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 2 to nucleotide 556; the nucleotide sequence of SEQ ID NO:14 from

nucleotide 53 to nucleotide 556; the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 367; the nucleotide sequence of the full-length protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:14.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 87 to amino acid 96 of SEQ ID NO:15; and

(c) the amino acid sequence encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 492 to nucleotide 602;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

(e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:18;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 492 to nucleotide 602; the nucleotide sequence of the full-length protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17, SEQ ID NO:16 or SEQ ID NO:19 .

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:18; and

(c) the amino acid sequence encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 154 to nucleotide 972;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 341;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:21;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 154 to nucleotide 972; the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 341; the nucleotide sequence of the full-length protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371; or the

nucleotide sequence of a mature protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 62.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 62;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 62.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 104 to nucleotide 892;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 299 to nucleotide 892;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 798 to nucleotide 1261;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:23;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:22 from nucleotide 104 to nucleotide 892; the nucleotide sequence of SEQ ID NO:22 from nucleotide 299 to nucleotide 892; the nucleotide sequence of SEQ ID NO:22 from nucleotide 798 to nucleotide 1261; the nucleotide sequence of the full-length protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371; or the nucleotide sequence of a mature protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23 from amino acid 53 to amino acid 89.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:22.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 53 to amino acid 89;

(c) fragments of the amino acid sequence of SEQ ID NO:23 comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:23 or the amino acid sequence of SEQ ID NO:23 from amino acid 53 to amino acid 89.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically

effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "bf171_6"

A polynucleotide of the present invention has been identified as clone "bf171_6". bf171_6 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bf171_6 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bf171_6 protein").

The nucleotide sequence of bf171_6 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bf171_6 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bf171_6 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for bf171_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bf171_6 demonstrated at least some similarity with sequences identified as AA147377 (zo39b08.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 589239 5'), AA190936 (zp83e01.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 626808 5'), AA287427 (zs52b05.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone), H77893 (ys09f08.r1 Homo sapiens cDNA), N72642 (yv74a12.r1 Homo sapiens cDNA clone), T25271 (Human gene signature HUMGS07433), T35346 (EST83197 Homo sapiens cDNA 5' end similar to None), and W27589 (34h1 Human retina cDNA randomly primed sublibrary Homo). Based upon sequence similarity, bf171_6 proteins and each similar protein or peptide may share at least some activity.

Clone "ck181_7"

A polynucleotide of the present invention has been identified as clone "ck181_7". ck181_7 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ck181_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ck181_7 protein").

The nucleotide sequence of ck181_7 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ck181_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ck181_7 should be approximately 1475 bp.

The nucleotide sequence disclosed herein for ck181_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. ck181_7 demonstrated at least some similarity with sequences identified as AA150370 (zl07e08.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491654 5'), H00151 (yl69h05.r1 Homo sapiens cDNA clone 43510 5'), N21123 (yx52f04.s1 Homo sapiens cDNA clone 265375 3'), N31138 (yx52f04.r1 Homo sapiens cDNA clone 265375 5'), R13827 (yf61h04.r1 Homo sapiens cDNA clone 26896 5' similar to SP:S42069 S42069 TEGT PROTEIN), and T19278 (Human gene signature HUMGS00295). The predicted amino acid sequence disclosed herein for ck181_7 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ck181_7 protein demonstrated at least some similarity to sequences identified as U88168 (weak similarity to rat TEGT protein (GI 456207) [Caenorhabditis elegans]). Based upon sequence similarity, ck181_7 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts seven potential transmembrane domains within the ck181_7 protein sequence, centered around amino acids 93, 136, 168, 206, 229, 258, and 283 of SEQ ID NO:4, respectively.

Clone "co736_3"

A polynucleotide of the present invention has been identified as clone "co736_3". co736_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. co736_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "co736_3 protein").

The nucleotide sequence of co736_3 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the co736_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 44 to 56 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 57, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone co736_3 should be approximately 1980 bp.

The nucleotide sequence disclosed herein for co736_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. co736_3 demonstrated at least some similarity with sequences

identified as H02676 (yj36g08.r1 Homo sapiens cDNA), H47499 (yp74c10.r1 Homo sapiens cDNA clone 193170 5'), Q53478 (MLL gene 8.3 kb BamHI genomic region), T91862 (yd54b07.s1 Homo sapiens cDNA clone 112021 3' similar to SP:LIN1_NYCCO P08548 LINE-1 REVERSE TRANSCRIPTASE ;contains Alu repetitive element;contains L1 repetitive element), U54776 (Human NTT gene, L1, Alu, and MER 38 repeat regions), Z73964 (Human DNA sequence from cosmid V698D2, between markers), and Z83843 (Human DNA sequence from PAC 368A4 on chromosome X. Contains ESTs, CELLULAR NUCLEIC ACID BINDING PROTEIN (CNBP) like gene and STSs). Based upon sequence similarity, co736_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the co736_3 protein sequence, one centered around amino acid 16 and another around amino acid 51 of SEQ ID NO:6. The nucleotide sequence of co736_3 indicates that it may contain one or more copies of the Alu repetitive element.

Clone "dm26_2"

A polynucleotide of the present invention has been identified as clone "dm26_2". dm26_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dm26_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dm26_2 protein").

The nucleotide sequence of dm26_2 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dm26_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 9 to 21 of SEQ ID NO:8 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 22, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dm26_2 should be approximately 3500 bp.

The nucleotide sequence disclosed herein for dm26_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dm26_2 demonstrated at least some similarity with sequences identified as AC000356 (Human cosmid g1346a312, complete sequence), F03454 (H.

sapiens partial cDNA sequence; clone c-1xh10), N42290 (yy06a07.r1 Homo sapiens cDNA clone 270420 5' similar to contains L1.t3 L1 repetitive element), N92463 (zb12e05.s1 Homo sapiens cDNA clone 301856 3'), N94118 (za25e06.r1 Homo sapiens cDNA clone 293602 5'), Q60160 (Human brain Expressed Sequence Tag EST02148), Z83745 (Human DNA sequence from PAC 453A3 contains EST and STS), and Z99129 (Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 425C14; HTGS phase 1. 1). The predicted amino acid sequence disclosed herein for dm26_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dm26_2 protein demonstrated at least some similarity to sequences identified as M22333 (unknown protein [Homo sapiens]), X61294 (L1 retroposon, a portion of its ORF2 sequence [Rattus norvegicus]), and Z81053 (E02A10.1 [Caenorhabditis elegans]). Based upon sequence similarity, dm26_2 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of dm26_2 indicates that it may contain one or more of the following repetitive elements: Alu, L1.

Clone "eq229_3"

A polynucleotide of the present invention has been identified as clone "eq229_3". eq229_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. eq229_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "eq229_3 protein").

The nucleotide sequence of the 5' portion of eq229_3 as presently determined is reported in SEQ ID NO:9. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:10. The predicted amino acid sequence of the eq229_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Amino acids 38 to 50 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 51, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of eq229_3, including the polyA tail, is reported in SEQ ID NO:11.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone eq229_3 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for eq229_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. eq229_3 demonstrated at least some similarity with sequences identified as N52034 (yz08g04.s1 Homo sapiens cDNA clone 282486 3') and W01791 (za72d06.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 298091 5'). Based upon sequence similarity, eq229_3 proteins and each similar protein or peptide may share at least some activity.

Clone "fh3_6"

A polynucleotide of the present invention has been identified as clone "fh3_6". fh3_6 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fh3_6 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fh3_6 protein").

The nucleotide sequence of fh3_6 as presently determined is reported in SEQ ID NO:12. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fh3_6 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:13. Amino acids 5 to 17 of SEQ ID NO:13 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 18. Another potential fh3_6 reading frame and predicted amino acid sequence is encoded by basepairs 765 to 1556 of SEQ ID NO:12 and is reported in SEQ ID NO:34. The overlapping open reading frames that encode SEQ ID NO:13 and SEQ ID NO:34 could be joined into a single oepn reading frame if a frameshift was introduced into the nucleotide sequence of SEQ ID NO:12 between base pairs 765 and 882.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fh3_6 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for fh3_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fh3_6 demonstrated at least some similarity with sequences identified as AA103102 (mo17f02.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus cDNA clone 553851 5'), W72947 (zd62g11.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 345284 3'), W74413 (zd62g11.r1 Soares fetal heart NbHH19W

Homo sapiens cDNA clone 345284 5'), and W88819 (zh71d11.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417525 5'). The predicted amino acid sequence disclosed herein for fh3_6 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fh3_6 protein demonstrated at least some similarity to sequences identified as Z81052) D2023.6 [Caenorhabditis elegans]). Based upon sequence similarity, fh3_6 proteins and each similar protein or peptide may share at least some activity. The Motifs computer progras predicts a prenyl group binding site (CAAX box) at amino acid 268 of SEQ ID NO:13.

Clone "fs87_3"

A polynucleotide of the present invention has been identified as clone "fs87_3". fs87_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fs87_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fs87_3 protein").

The nucleotide sequence of fs87_3 as presently determined is reported in SEQ ID NO:14. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fs87_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15. Amino acids 5 to 17 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 18, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fs87_3 should be approximately 1300 bp.

The nucleotide sequence disclosed herein for fs87_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fs87_3 demonstrated at least some similarity with sequences identified as AA223699 (zr10c04.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 651078 3') and AA287263 (zs49h08.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:700863 5' similar to SW:CC91_YEAST P41733 CELL DIVISION CONTROL PROTEIN 91). The predicted amino acid sequence disclosed herein for fs87_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fs87_3 protein demonstrated at least

some similarity to sequences identified as L31649 (cdc91 [Saccharomyces cerevisiae]), S72417 (E2 {patient 3} [hepatitis C virus]), U06711 (tracheobronchial mucin [Homo sapiens]), Z75550 (T22C1.3 [Caenorhabditis elegans]), and Z98598 (hypothetical protein [Schizosaccharomyces pombe]). Based upon sequence similarity, fs87_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two additional potential transmembrane domains within the fs87_3 protein sequence, one centered around amino acid 90 and another around amino acid 170 of SEQ ID NO:15.

Clone "fy530_2"

A polynucleotide of the present invention has been identified as clone "fy530_2". fy530_2 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fy530_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fy530_2 protein").

The nucleotide sequence of the 5' portion of fy530_2 as presently determined is reported in SEQ ID NO:16. An additional internal nucleotide sequence from fy530_2 as presently determined is reported in SEQ ID NO:17. What applicants believe is the proper reading frame and the predicted amino acid sequence encoded by such internal sequence is reported in SEQ ID NO:18. Additional nucleotide sequence from the 3' portion of fy530_2, including the polyA tail, is reported in SEQ ID NO:19.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fy530_2 should be approximately 3550 bp.

The nucleotide sequence disclosed herein for fy530_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fy530_2 demonstrated at least some similarity with sequences identified as AA029852 (zk11b04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470191 3'), AA118938 (mp64g01.r1 Soares 2NbMT Mus musculus cDNA clone 574032 5'), L39210 (Human inosine monophosphate dehydrogenase type II gene, complete cds), N51229 (yz13b07.s1 Homo sapiens cDNA clone 282901 3'), and X95808 (H.sapiens mRNA for protein encoded by a candidate gene, DXS6673E, for mental retardation). The predicted amino acid sequence disclosed herein for fy530_2 was searched against the

GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fy530_2 protein demonstrated at least some similarity to sequences identified as X95808 (X-linked mental retardation candidate gene [Homo sapiens]). Based upon sequence similarity, fy530_2 proteins and each similar protein or peptide may share at least some activity.

Clone "ge51_1"

A polynucleotide of the present invention has been identified as clone "ge51_1". ge51_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ge51_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ge51_1 protein").

The nucleotide sequence of ge51_1 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ge51_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ge51_1 should be approximately 1850 bp.

The nucleotide sequence disclosed herein for ge51_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ge51_1 demonstrated at least some similarity with sequences identified as AA219716 (zq98d02.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 650019 5'), AA434286 (zw30f01.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 770809 5' similar to SW:NALS_BOVIN P08037 N-ACETYLLACTOSAMINE SYNTHASE), D61576 (Human fetal brain cDNA 5'-end GEN-419H03), H30715 (yo78h01.r1 Homo sapiens cDNA clone 184081 5'), T80315 (yd07b08.r1 Homo sapiens cDNA clone 24966 5'), U19889 (Gallus gallus beta-1,4-galactosyltransferase (CKII) mRNA, complete cds), and W90417 (zh72h01.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417649 3'). The predicted amino acid sequence disclosed herein for ge51_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ge51_1 protein demonstrated at least some similarity to sequences identified as

M70433 (beta-1,4-galactosyltransferase [Homo sapiens]), R05932 (Human beta-1,4-galactosyltransferase), and beta-1,4-galactosyltransferases from several other species. Based upon sequence similarity, ge51_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the ge51_1 protein sequence, one centered around amino acid X20 and another around amino acid 90 of SEQ ID NO:21.

Clone "gx183_1"

A polynucleotide of the present invention has been identified as clone "gx183_1". gx183_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gx183_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gx183_1 protein").

The nucleotide sequence of gx183_1 as presently determined is reported in SEQ ID NO:22. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gx183_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:23. Amino acids 53 to 65 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 66, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gx183_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for gx183_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gx183_1 demonstrated at least some similarity with sequences identified as AA010474 (zi09a06.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 430258 5'), H01847 (yj28f09.r1 Homo sapiens cDNA clone 150089 5'), L38971 (Mus musculus (E25) mRNA, complete cds), Q60909 (Human brain Expressed Sequence Tag EST00998), W37875 zc13c01.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322176 3'), and W72197 (zd69e11.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 345932 3'). The predicted amino acid sequence disclosed herein for gx183_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted gx183_1 protein demonstrated at least

some similarity to sequences identified as AL021786 (dJ696H22.1 (mouse E25 like protein) [Homo sapiens]) and L38971 (putative [Mus musculus]). Based upon sequence similarity, gx183_1 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones bf171_6, ck181_7, co736_3, dm26_2, eq229_3, fh3_6, fs87_3, fy530_2, ge51_1, and gx183_1 were deposited on March 25, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98371, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the

oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
bf171_6	SEQ ID NO:24
ck181_7	SEQ ID NO:25
co736_3	SEQ ID NO:26
dm26_2	SEQ ID NO:27
eq229_3	SEQ ID NO:28
fh3_6	SEQ ID NO:29
fs87_3	SEQ ID NO:30
fy530_2	SEQ ID NO:31
ge51_1	SEQ ID NO:32
gx183_1	SEQ ID NO:33

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to

the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided.

Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that

shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus*, for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* **22**: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* **25**: 682-690; Lyons *et al.*, 1997, *Nature Genetics* **15**: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* **13(10)**: 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90%

identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1521 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTAACCTTCT TCTGCGCGGC TGCAGCTCGG GACTTCGGCC TGACCCAGCC CCCATGGCTT 60

CAGAAGAGCT ACAGAAAGAT CTAGAAGAGG TAAAGGTGTT GCTGGAAAAG GCTACTAGGA 120

AAAGAGTACG TGATGCCCTT ACAGCTGAAA AATCCAAGAT TGAGACAGAA ATCAAGAACA 180

AGATGCAACA GAAATCACAG AAGAAAGCAG AACTTCTTGA TAATGAAAAA CCAGCTGCTG 240

TGGTTGCTCC CATAACAACG GGCTATACGG TGAAAATCAG TAATTATGGA TGGGATCAGT 300

CAGATAAGTT TGTGAAAATC TACATTACCT TAACTGGAGT TCATCAAGTT CCCACTGAGA 360

ATGTGCAGGT GCATTTCACA GAGAGGTCAT TTGATCTTTT GGTAAAGAAT CTAAATGGGA 420

AGAGTTACTC CATGATTGTG AACAATCTCT TGAAACCCAT CTCTGTGGAA GGCAGTTCAA 480

AAAAAGTCAA GACTGATACA GTTCTTATAT TGTGTAGAAA GAAAGTGGAA AACACAAGGT 540

GGGATTACCT GACCCAGGTT GAAAAGGAGT GCAAAGAAAA AGAGAAGCCC TCCTATGACA 600

CTGAAACAGA TCCTAGTGAG GGATTGATGA ATGTTCTAAA GAAAATTTAT GAAGATGGAG 660

ACGATGATAT GAAGCGAACC ATTAATAAAG CCTGGGTGGA ATCAAGAGAG AAGCAAGCCA 720

AAGGAGACAC GGAATTTTGA GACTTTAAAG TCGTTTTGGG AACTGTGATG TGATGTGGAA 780

ATACTGATGT TTCCAGTAAG GGAATATTGG TGAGCTGCAT ATATAAATTT GACAGATAGC 840

TATTTACATA GCCTTCTAAG TAAAGGCAAT GAATTCTCCA TTTCCTACTG GAGGATTTAT 900

TTAAATAAAA TATGCTTATT AAACACTCCT GCAAAGATGG TTTTATTAGT ACCCTGGTCA 960

TTTTGTTCAA GGAAGGGTTA TATTGCATTC TCACGTGAAA TATAAAAAGC AAGTCTTGCC 1020

CAATAAAAAC GCTACATTGT GTGTATTTTT TGTTCAGCTA AGAATTGGAA AAGTATTTGC 1080

TTGCCTTTTA AGTTACTGAC ATCAGCTTCC ACCAGTGTAA AAATTGAGTA AAACCTGAAG 1140

TTTTGCATAA AATGCAAATC GGTGCCTGTG CTTGAAGGTT GCTGTAGAGC ATCTGACCCC 1200

TTATTACCAC CTTAAGCAAT GTATATGCCA TGCATTACCA TGCACTAATT CAATCACAGG 1260

TGTTTCTATC TAGATTTAAA TATATTTGTC AATGAATGTG GAATAGAAAA TCTAAACATG 1320

ACAATAATAG ACATATCTTT GTATGGTACC AGTTAGTTTT GCCGTGGATC AGATGGTTTA 1380

TAAAAGTAAT AACCATAAAG CAAAAAATAA TTTGAAAGCC CGTCTATTCC TATGCTCAAT 1440

AAAGTTAAGT TTTTTTTCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1500

AAAAAAAAAA AAAAAAAAAA A 1521

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Ser Glu Glu Leu Gln Lys Asp Leu Glu Glu Val Lys Val Leu
1 5 10 15

Leu Glu Lys Ala Thr Arg Lys Arg Val Arg Asp Ala Leu Thr Ala Glu
20 25 30

Lys Ser Lys Ile Glu Thr Glu Ile Lys Asn Lys Met Gln Gln Lys Ser
35 40 45

Gln Lys Lys Ala Glu Leu Leu Asp Asn Glu Lys Pro Ala Ala Val Val
50 55 60

Ala Pro Ile Thr Thr Gly Tyr Thr Val Lys Ile Ser Asn Tyr Gly Trp
65 70 75 80

Asp Gln Ser Asp Lys Phe Val Lys Ile Tyr Ile Thr Leu Thr Gly Val
85 90 95

His Gln Val Pro Thr Glu Asn Val Gln Val His Phe Thr Glu Arg Ser
100 105 110

Phe Asp Leu Leu Val Lys Asn Leu Asn Gly Lys Ser Tyr Ser Met Ile
115 120 125

Val Asn Asn Leu Leu Lys Pro Ile Ser Val Glu Gly Ser Ser Lys Lys
130 135 140

Val Lys Thr Asp Thr Val Leu Ile Leu Cys Arg Lys Lys Val Glu Asn
145 150 155 160

Thr Arg Trp Asp Tyr Leu Thr Gln Val Glu Lys Glu Cys Lys Glu Lys
165 170 175

Glu Lys Pro Ser Tyr Asp Thr Glu Thr Asp Pro Ser Glu Gly Leu Met
180 185 190

Asn Val Leu Lys Lys Ile Tyr Glu Asp Gly Asp Asp Asp Met Lys Arg
195 200 205

Thr Ile Asn Lys Ala Trp Val Glu Ser Arg Glu Lys Gln Ala Lys Gly
210 215 220

Asp Thr Glu Phe
225

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1394 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGCGTCATGC AGTGCGCCGG AGGAACTGTG CTCTTTGAGG CCGACGCTAG GGGCCCGGAA 60

GGGAAACTGC GAGGCGAAGG TGACCGGGGA CCGAGCATTT CAGATCTGCT CGGTAGACCT 120

GGTGCACCAC CACCATGTTG GCTGCAAGGC TGGTGTGTCT CCGGACACTA CCTTCTAGGG 180

TTTTCCACCC AGCTTTCACC AAGGCCTCCC CTGTTGTGAA GAATTCCATC ACGAAGAATC 240

AATGGCTGTT AACACCTAGC AGGGAATATG CCACCAAAAC AAGAATTGGG ATCCGGCGTG 300

GGAGAACTGG CCAAGAACTC AAAGAGGCAG CATTGGAACC ATCGATGGAA AAAATATTTA 360

AAATTGATCA GATGGGAAGA TGGTTTGTTG CTGGAGGGGC TGCTGTTGGT CTTGGAGCAT 420

TGTGCTACTA TGGCTTGGGA CTGTCTAATG AGATTGGAGC TATTGAAAAG GCTGTAATTT 480

GGCCTCAGTA TGTCAAGGAT AGAATTCATT CCACCTATAT GTACTTAGCA GGGAGTATTG 540

GTTTAACAGC TTTGTCTGCC ATAGCAATCA GCAGAACGCC TGTTCTCATG AACTTCATGA 600

TGAGAGGCTC TTGGGTGACA ATTGGTGTGA CCTTTGCAGC CATGGTTGGA GCTGGAATGC 660

TGGTACGATC AATACCATAT GACCAGAGCC CAGGCCCAAA GCATCTTGCT TGGTTGCTAC 720

ATTCTGGTGT GATGGGTGCA GTGGTGGCTC CTCTGACAAT ATTAGGGGGT CCTCTTCTCA 780

TCAGAGCTGC ATGGTACACA GCTGGCATTG TGGGAGGCCT CTCCACTGTG GCCATGTGTG 840

CGCCCAGTGA AAAGTTTCTG AACATGGGTG CACCCCTGGG AGTGGGCCTG GGTCTCGTCT 900

TTGTGTCCTC ATTGGGATCT ATGTTTCTTC CACCTACCAC CGTGGCTGGT GCCACTCTTT 960

ACTCAGTGGC AATGTACGGT GGATTAGTTC TTTTCAGCAT GTTCCTTCTG TATGATACCC 1020

AGAAAGTAAT CAAGCGTGCA GAAGTATCAC CAATGTATGG AGTTCAAAAA TATGATCCCA 1080

TTAACTCGAT GCTGAGTATC TACATGGATA CATTAAATAT ATTTATGCGA GTTGCAACTA 1140

TGCTGGCAAC TGGAGGCAAC AGAAAGAAAT GAAGTGACTC AGCTTCTGGC TTCTCTGCTA 1200

CATCAAATAT CTTGTTTAAT GGGGCAGATA TGCATTAAAT AGTTTGTACA AGCAGCTTTC 1260

GTTGAAGTTT AGAAGATAAG AAACATGTCA TCATATTTAA ATGTTCCGGT AATGTGATGC 1320

CTCAGGTCTG CCTTTTTTTC TGGAGAATAA ATGCAGTAAT CCTCTCCCAA ATAAGCACAA 1380

AAAAAAAAAA AAAA 1394

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 345 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Leu Ala Ala Arg Leu Val Cys Leu Arg Thr Leu Pro Ser Arg Val
1 5 10 15

Phe His Pro Ala Phe Thr Lys Ala Ser Pro Val Val Lys Asn Ser Ile
20 25 30

Thr Lys Asn Gln Trp Leu Leu Thr Pro Ser Arg Glu Tyr Ala Thr Lys
35 40 45

Thr Arg Ile Gly Ile Arg Arg Gly Arg Thr Gly Gln Glu Leu Lys Glu
50 55 60

Ala Ala Leu Glu Pro Ser Met Glu Lys Ile Phe Lys Ile Asp Gln Met
65 70 75 80

Gly Arg Trp Phe Val Ala Gly Gly Ala Ala Val Gly Leu Gly Ala Leu
85 90 95

Cys Tyr Tyr Gly Leu Gly Leu Ser Asn Glu Ile Gly Ala Ile Glu Lys
100 105 110

Ala Val Ile Trp Pro Gln Tyr Val Lys Asp Arg Ile His Ser Thr Tyr
115 120 125

```
Met Tyr Leu Ala Gly Ser Ile Gly Leu Thr Ala Leu Ser Ala Ile Ala
    130                 135                 140

Ile Ser Arg Thr Pro Val Leu Met Asn Phe Met Met Arg Gly Ser Trp
145                 150                 155                 160

Val Thr Ile Gly Val Thr Phe Ala Ala Met Val Gly Ala Gly Met Leu
                165                 170                 175

Val Arg Ser Ile Pro Tyr Asp Gln Ser Pro Gly Pro Lys His Leu Ala
            180                 185                 190

Trp Leu Leu His Ser Gly Val Met Gly Ala Val Val Ala Pro Leu Thr
        195                 200                 205

Ile Leu Gly Gly Pro Leu Leu Ile Arg Ala Ala Trp Tyr Thr Ala Gly
    210                 215                 220

Ile Val Gly Gly Leu Ser Thr Val Ala Met Cys Ala Pro Ser Glu Lys
225                 230                 235                 240

Phe Leu Asn Met Gly Ala Pro Leu Gly Val Gly Leu Gly Leu Val Phe
                245                 250                 255

Val Ser Ser Leu Gly Ser Met Phe Leu Pro Pro Thr Thr Val Ala Gly
            260                 265                 270

Ala Thr Leu Tyr Ser Val Ala Met Tyr Gly Gly Leu Val Leu Phe Ser
        275                 280                 285

Met Phe Leu Leu Tyr Asp Thr Gln Lys Val Ile Lys Arg Ala Glu Val
    290                 295                 300

Ser Pro Met Tyr Gly Val Gln Lys Tyr Asp Pro Ile Asn Ser Met Leu
305                 310                 315                 320

Ser Ile Tyr Met Asp Thr Leu Asn Ile Phe Met Arg Val Ala Thr Met
                325                 330                 335

Leu Ala Thr Gly Gly Asn Arg Lys Lys
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1908 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCTTTTTTTT TTTTTTTTGG TTGAGATGGG GTCTCGCCAT GTTTCCCACA CTGATCTTGA 60

ACTCCTGGGC TCCAGGAATT CTCCTACTTT GGCCTCCCAA AGTGTTGGGA ATATTGGCAT 120

GAACCACAGC ACCTGACTTG CATATTTGTG AATTCCCCAA ATTGCTTTTT TTAAATTGAT 180

TTCTAATTTC ATTTCATTGT TATGGGGAAC ATACTTTGTA TGGTTTCAAT GTTTTAAAAT 240

TAATTGAGAC TTGTTTTATG ACTTAGCATA TGGTCTGTGT TGAAGAAGGC TCCATGTACA 300

CTTGAGAATA ATATGTATAC TGTGGTTGTT GGGTGGATTT TCTATGTATG TTTARGTGAT 360

ATGGTTTTAT AGTGTTGTTT AARTCTTCTA TTTTCTTCTT TTTCTGCCCA GTTTTATTTT 420

TGAAAGCATA CTGARGTCTC CAACTCARTG CCTTAGCCTC CTGAGCAGTT GGGACTACAG 480

GCATACGCCA CTACACCCAG CAATTTTTTT GTATTTTTCT GTAGAGACAG AGTTTCACCA 540

TGTTGCCTAG GCTGGTCTCA GATTCCTGGA CTCAAGTGAT CTCGATTCCC GGCCTCTGCC 600

TCCCGGGGTG CTGGGATTGC AGGCATGAGC TACTATGCCT GGCAAATTTT ATTTTTCCTT 660

TTATTTTGTC ACATAATTAA AGCTACTCCA GAATTCCCTT GATTTCTGCT TGCCTGGTAT 720

ATCTTTTTTC CATTTTTTAA CTGTCAGCCT TTTTTGTGCC TGTTAATCTA AAGTATGTGT 780

TTCGTAGATA ATATGTAGCT GGATCATATT TTAAAAATAT TTATTCTGCC AAGCTCTGTC 840

TTTTGATTGG AGTATTCTTT CATTTATGTT TGTAATTACT GATGAGGGGG GCACTAATGT 900

CTGCTGTTTT GCTATTTGTT TCCCCATGTC TTATGTCTTC ATTACTGACT TTTTTATTAA 960

ACAACTATTT TCTTGGGTAC CATTTTAAGT CCCTCTCCCA CTCATTTTTT AATGTTTTTT 1020

TGTGTTTACT TTTGTTTTTA TTGTTTGCCC TGATATTAAA ATTAACATTT TACCTTGAAA 1080

TAGTTGGCTT CAGATTAATA TCAACTTAGT TTCAATAGCA TAGGAAATTT GCTTCACTAT 1140

ATTTCCATTT TCTCCCCGTC CTTTGTGCTA TTATTACTAT ACCAATTAGA TCTCTACACA 1200

ATATAGGCAT ATCAACACAT TTTGTAATTA TTTCCTTATC CAGTTGTCTT TTAATATAGA 1260

TCTGTGAAGA AAAGTATTAC ACAAATAGAT CTATTCTGTT TTTTATAATT ATTTAATTAC 1320

CTTTGGTGGT GCTGTTTATT TTTCATGCAT TTGAGTTACT GTCTAGTATT CATTCATTTC 1380

TCTCTGAATC ACTCCCTTTA GTATTGCTTG TAGGGCAGGT CTGCTAGCAT TGAATTCTTT 1440

TAATTTTTGT GACTCTGCAA ATGCCATAAT TTCTCTTTTG TTTGTGAAGG ATAGTTTTGC 1500

TAGATACAGA ATTTGCAGTT GGCATTCTTT TTACTTTAGC AGTTTGAAAA TATTTCCCAT 1560

TGTTGGCCGG GCACAGTGGC TCACGCCTGT GGTCCTAGCA CTTTGGGAGG CCGAGGCGGG 1620

CGGATCGTCT GGGGTCGGGA GTTCGGGACC GGCCTGGCCA ATATGGTGAG GCCCTGTTTC 1680

TGCTAAAATA TAAAAATTGG CTGGGCATGA TGGCGGGTGC CTCTAGTCCC AGCTGCTCGG 1740

GAGGCTGAGG TGGGGGAGTC GCTTGAGCCC GGGAGATGAT GGCTGTGGTG AGCCGGGATG 1800

GCGCCGCTGC ACTCCGGCCT GGGCGGCTGA GTGAGACTCC ATCCCCGAAA AAAAAAAAAA 1860

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA 1908

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Arg Gly Ala Leu Met Ser Ala Val Leu Leu Phe Val Ser Pro Cys
1 5 10 15

Leu Met Ser Ser Leu Leu Thr Phe Leu Leu Asn Asn Tyr Phe Leu Gly
20 25 30

Tyr His Phe Lys Ser Leu Ser His Ser Phe Phe Asn Val Phe Leu Cys
35 40 45

Leu Leu Leu Phe Leu Leu Phe Ala Leu Ile Leu Lys Leu Thr Phe Tyr
50 55 60

Leu Glu Ile Val Gly Phe Arg Leu Ile Ser Thr
65 70 75

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3076 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTTTTTTTT TTTTTTCAAT TTCATTTAGT TCTGCCCTGA TCTTGGTTAT TTCCTTTTTT 60

TCTGCTGGGT TTGGGTTTGG TTTGTTCTTA TTTCTCTAGT TCCTTGAGGT GTGACCTTAR 120

AATGTCAATT TGTGCTCTTT CAATCTTTTT GATGTAGGCG TTGAGGGCTG TGGACTTTTC 180

TCTTGGCACT CCCTTTGGTG TATCCCARAG GTTTTGATAG GTTGTGTCAT TATTGCAATT 240

CAGTTTGAAG AATTTCTTAA TCTCCACCTT GATTTTGTTT TTGACCCAAT GCTCATTCAG 300

GAGCAGGTTA TTTACTTTCC ATGTACTTGC ATGGCTTTGA AGCTTCCTTT TGGAGTTGAT 360

TTCCAGTTTT ATTCCACTGT GATTTGAGAG AGTGCTTTAC ATAATTTCAA TTTTCTTAAT 420

TTTATTAAGG CTCGTTTTAT GGCCTATAAT ATGGTCTATC TTGGAGAAAG TTCCATGCAC 480

TGTAGAATAG AATGTGTATT CTGTGGTTGT TGGATGAAAT GTTCTGCATA TATTCCTAGA 540

TTGCCTCCCC ACAAAAGGTT GCATCAATGT CTGTGTTTCT CTACACCATC TCACCCTTGC 600

CAACTTCGGG TTTCATCAGA CCTTACTGAT TGTCAGTATG ATCTGTGAAA CAAATCTCTC 660

AGTTTTGATT TGCATTTTTT AAATTATGAG AGCTTGAACA CCATTTTACA TGTTTATTGG 720

CTGTTGTTAT TTCCTTTTTG AGATCTGTTC GTTATATGCT TTGCCCGTTT TTCTGTTGGG 780

TGGTTATTAT TTTTCTTATT GAATGGTATA AGCTCTTTGT AAGTTAAGGA CATTAGCCCT 840

TAGTCAGATA TTTTGACTTA GGTTTTAATT TTTTTCCACA CAGAAGTTTT AAGCTCTGTG 900

GCAAATTTAT CAGTCTTATA TCACTACAGG GTTATAAATA TTAGYTATCA CTTCGGGTTT 960

GTGTCTTGCT TAGAAAGCMT CATTTGAAGA TTGTAAATGT TAGTAAGTTT CCCCATATTT 1020

TCCTCTAGGA CTTCCATGGT TTAATTTGTT TTGTTTAAAY TAGGAATTGG CATTCACATC 1080

CTYTTTTGTC CCAGGTCTCA GAGGTCCCTT GTATCTTATA GAGCAGTATT GTTTTATGTT 1140

ATTTTCCCAT GTATAATTTA AAAACAAAAT ACGTTGTTCA AAACAAAATA CAGTGGCAGC 1200

AGATAATGGC AGTATCTCTG TAACTGCTGG TAAACTGTAT TTCATAGTGA AGTGTTCATA 1260

AACTAAAGAG TCATTGATTT GGTTTCCTGG CTAATTAAAA TCTGAATTCC ATTTGAAGTT 1320

CCATTGAAAT CATGGTTTTA CTCTATAGCA GTGGATGTTT TTTCCCAACC TTTCTGATAT 1380

TTTTTTCCTT CCTGAGACAG GGTCTTGCTC TGTCACCTGG GATGGAGTGT AGTTGCACCA 1440

TCAAGGCTTA CTGCAGTCTC AACTCTCTGA GCTCAAGTGA TCCTGCCACC TCAGCCTCTT 1500

GAGTAGCAAG GATTACAGGC ACCTACCACT ATGCCTGGCT AATTTTTATA TTTTTTGTAG 1560

AGATGGATTC TCACTATGTT GCCCGGGCTC ATCTTGAACT CGAGCTCAAG CAATCTGTCC 1620

ATCTTGGCCT CCCAAAGTGC TGGGATTATA GGCGTGAGCC ACTGCACCTG GCCCCTTTCT 1680

GATTATTTTA ATCTATCTTT AAATGTTCAA AGTGATTTGC CTAATTCATT TAAAGCATAT 1740

TTAGTTTTTT TTAAATTGAG TGTATTTTAT CTAGATATTT TTAAAAGGCA GCATCTAACC 1800

TTGGATTTTA TAAATACATC TAAATTTGTT ATTTCCAGAA TGCTTCAAAA CAGATCTCTG 1860

TAGCCTCGTG CTTTGTTATT GTTAGGTTTT TTTTTTTTGT TTTGAGACAG GGTCTTGCTC 1920

TATCTGGAGT GCAGTGGCAC AGTCATAGCT CACTGTACCC TCAAACTCCT AAACTCAAGT 1980

AATCCTCCCA TCTCAGCCTC CTGAGTAGTT GGGACCACAG TCATGCACCA GCATGCCTGG 2040

CTAATTTTTT AAATTTTGTT CTTAATAGAG ACAGAGTCTT GCTGTGTTGT TCAGGCTGGT 2100

CTCAAACTCC TGGGCTCAAG CGATCCTCCC ACCTCAGCCT CCTAAAGTGC TGAGATTACG 2160

GATGTGAATC ATTACACCCA GCCTATTAAT GGTTTTGTAT AGCAAGTCTT TTGTGGGTGG 2220

TGGAAAGATG AAGTGCTGTG AAATATTGTA GGAGCAGAAA CTTGAAATGT GGCAAAAACC 2280

ACATGGGCAA AATTTCTGTC TCTTTTCTTA TTTTTGCTTT TTTGTTTAAA GGTTTTTCTA 2340

TTGGGAAAGC TACTGATCGG ATGGATGCTT TCAGGAAAGC AAAGAACAGA GCAGTTCACC 2400

ATTTGCATTA TATAGAACGA TATGAAGACC ATACAATATT CCATGATATT TCATTAAGAT 2460

TTAAAAGGAC GCATATCAAG ATGAAGAAAC AACCCAAAGG TTACGGCCTC CGCTGCCACA 2520

GGGCCATCAT CACCATCTGC CGGCTCATTG GCATCAAAGA CATGTATGCC AAGGTCTCTG 2580

GGTCCATTAA TATGCTCAGC CTCACCCAGG GCCTCTTCCG TGGGCTCTCC AGACAGGAAA 2640

CCCATCAACA GCTGGCTGAT AAGAAGGGCC TCCATGTTGT GGAAATCCGG GAGGAATGTG 2700

GCCCTCTGCC CATTGTGGTT GCGTCCCCCC GGGGGCCCTT GAGGAAGGAT CCAGAGCCAG 2760

AAGATGAGGT TCCAGACGTC AAACTGGACT GGGAAGATGT GAAGACTGCA CAGGGAATGA 2820

AGCGCTCTGT GTGGTCTAAT TTGAAGAGAG CCGCCACGTA ACCTCTCTGG CCTTGTGCAG 2880

CCAGTTCCTG TGCTGCCCTG CACCTAGGAG AGACTCAGCC CCTCACAGCT TGGGATGTTA 2940

CCTTGCCTTT TGTTTGTTTT GAGGGAAGTT TAATCTTTAA ACTCTTTGGA AATAAATAAT 3000

TATAGCTTTC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3060

AAAAAAAAAA AAAAAA 3076

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 192 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Gly Lys Ile Ser Val Ser Phe Leu Ile Phe Ala Phe Leu Phe Lys
1 5 10 15

Gly Phe Ser Ile Gly Lys Ala Thr Asp Arg Met Asp Ala Phe Arg Lys
20 25 30

Ala Lys Asn Arg Ala Val His His Leu His Tyr Ile Glu Arg Tyr Glu
35 40 45

Asp His Thr Ile Phe His Asp Ile Ser Leu Arg Phe Lys Arg Thr His
50 55 60

Ile Lys Met Lys Lys Gln Pro Lys Gly Tyr Gly Leu Arg Cys His Arg
65 70 75 80

Ala Ile Ile Thr Ile Cys Arg Leu Ile Gly Ile Lys Asp Met Tyr Ala
85 90 95

Lys Val Ser Gly Ser Ile Asn Met Leu Ser Leu Thr Gln Gly Leu Phe
100 105 110

Arg Gly Leu Ser Arg Gln Glu Thr His Gln Gln Leu Ala Asp Lys Lys
115 120 125

Gly Leu His Val Val Glu Ile Arg Glu Glu Cys Gly Pro Leu Pro Ile
130 135 140

Val Val Ala Ser Pro Arg Gly Pro Leu Arg Lys Asp Pro Glu Pro Glu
145 150 155 160

Asp Glu Val Pro Asp Val Lys Leu Asp Trp Glu Asp Val Lys Thr Ala
165 170 175

Gln Gly Met Lys Arg Ser Val Trp Ser Asn Leu Lys Arg Ala Ala Thr
180 185 190

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 683 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCCAAGTGC GCATGGGGAC GCTATAGCAA TTCGTTTGCT GTCCTTCCTC TCCTTCGAAG 60

ATGACAAGGC CTACCATCGT TTCTTCCTGC CTTTGGGCCG TCAGGCAGTT GGTTGGGACC 120

CGCTCCAACC CTCGGTTCTT CCTGCAATAC AGTGGATACA ATTTGTCATG GCTACTCTGA 180

GTGTTATAGG TTCAAGTTCA CTTATTGCCT ATGCTGTATT CCATAATATA CAGAAATCTC 240

CAGAGATAAG ACCACTTTTT TATCTGAGCT TCTGTGACCT GCTCCTGGGA CTTTGCTGGC 300

TCACGGAGAC ACTTCTCTAT GGAGCTTCAG TAGCAAATAA GGACATCATC TGCTATAACC 360

TACAAGCAGT TGGACAGATA TTCTACATTT CCTCATTTCT CTACACCGTC AATTACATCT 420

GGTATTTGTA CACAGAGCTG AGGATGAAAC ACACCCAGAG TGGACAGAGC ACATCTCCAC 480

TGGTGATAGA TTATACTTGT CGAGTTGGTC AAATGGCCTT TGTTTTCTCA AGCCTGATAC 540

CTCTGCTATT GATGACACCT GTATTCTGTC TGGGAAATAC TAGTGAATGT TTCCAAAACT 600

TCAGTCAGAG CCACAAGTGT ATCTTGATGC ACTCACCACC ATCAGCCATG GCTGAACTTC 660

CACCTTCTGC CAACACATCT GTC 683

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 172 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Thr Leu Ser Val Ile Gly Ser Ser Ser Leu Ile Ala Tyr Ala
1 5 10 15

Val Phe His Asn Ile Gln Lys Ser Pro Glu Ile Arg Pro Leu Phe Tyr
20 25 30

Leu Ser Phe Cys Asp Leu Leu Leu Gly Leu Cys Trp Leu Thr Glu Thr
35 40 45

Leu Leu Tyr Gly Ala Ser Val Ala Asn Lys Asp Ile Ile Cys Tyr Asn
50 55 60

Leu Gln Ala Val Gly Gln Ile Phe Tyr Ile Ser Ser Phe Leu Tyr Thr

65 70 75 80

Val Asn Tyr Ile Trp Tyr Leu Tyr Thr Glu Leu Arg Met Lys His Thr
85 90 95

Gln Ser Gly Gln Ser Thr Ser Pro Leu Val Ile Asp Tyr Thr Cys Arg
100 105 110

Val Gly Gln Met Ala Phe Val Phe Ser Ser Leu Ile Pro Leu Leu Leu
115 120 125

Met Thr Pro Val Phe Cys Leu Gly Asn Thr Ser Glu Cys Phe Gln Asn
130 135 140

Phe Ser Gln Ser His Lys Cys Ile Leu Met His Ser Pro Pro Ser Ala
145 150 155 160

Met Ala Glu Leu Pro Pro Ser Ala Asn Thr Ser Val
165 170

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 524 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATATGGCTGG ACGCAGCACA AATTCCACCA ACTAAAGCAG GAGGCTCGGC GTGATGCAGA 60

TACCCAGACA CCATTATTAT GCTCACAGAA GAGATTCTAT AGCAGGGGCT TAAATTCACT 120

GGAATCCACC CTGACTTTTC CTGCCAGTAC TTCTACCATT TTTTGAAACT ACAATACTGG 180

AACATCCAGG AACTGGAGTT ATTCTACGCT AATGGATTGG AAAGAATGTT GGGAAAGGAC 240

ATCTTAAATC TTTTCTAACT ATGCCCTAAA CTGCAGAACT CAAAGGAAAT ATAGTGCCAT 300

TGTTAGTAGT CATTCTAGAT GAATTGGGAG TATCTCTCCA GTTATTCCCA GATTCACTAG 360

TGATCCTTAA AGTCTCTATT CAGGGAGAGG AAGACACTTT CCATCTCAGA GATAGACTCG 420

TGTTACCTTG ATGGATATTG GATTTGTCTA AGTCTCTTCT AGAAAAAATA AATTCTAGAT 480

TATTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA 524

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2171 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCCCGCTACC GGGTTGCGGC CGGAAGCCGG GCGCCGCGGC TCTGCTTCCC TCGGGGATCT 60

GGCGACATGG CCAGAAAGGC TCTCAAGCTT GCTTCGTGGA CCAGCATGGC TCTTGCTGCC 120

TCTGGCATCT ACTTCTACAG TAACAAGTAC TTGGACCCTA ATGACTTTGG CGCTGTCAGG 180

GTGGGCAGAG CAGTTGCTAC GACGGCTGTC ATCAGTTACG ACTACCTCAC TTCCCTGAAG 240

AGTGTCCCTT ATGGCTCAGA GGAGTACTTG CAGCTGAGAT CTAAGGTGCA CCTTCGCTCT 300

GCCAGGCGTC TCTGTGAGCT CTGCTGTGCC AACCGGGGCA CCTTCATCAA GGTGGGCCAG 360

CACCTGGGGG CTCTGGACTA CCTGTTGCCA GAGGAGTACA CCAGCACGCT GAAGGTACTG 420

CACAGCCAGG CTCCACAGAG CAGCATGCAA GAGATCCGCC AGGTCATCCG AGAAGATCTG 480

GGCAAGGAGG TGCTCGTTCT GGCTGTGAAG CAGCTGTTCC CAGAGTTTGA GTTTATGTGG 540

CTTGTGGATG AAGCCAAGAA GAACCTGCCT TTGGAGCTGG ATTTCCTCAA TGAAGGGAGG 600

AATGCTGAGA AGGTGTCCCA GATGCTCAGG CATTTTGACT TCTTGAAGGT CCCCCGAATC 660

CACTGGGACC TGTCCACGGA GCGGGTCCTC CTGATGGAGT TTGTGGATGG CGGGCAGGTC 720

AATGACAGAG ACTACATGGA GAGGAACAAG ATCGACGTCA ATGAGGTGAG GTCAAGAGCT 780

CAGGGCTGCT GTGCCGGGGA ACGTGGGCTT GGTCAAGGCT GCCCAGGAAG TGCCTGTGTG 840

TCCAGATCTC ACGCCACCTG GGCAAGATGT ATAGTGAGAT GATCTTCGTC AATGGCTTCG 900

TGCACTGCGA TCCCCACCCC GGCAATGTAC TGGTGCGGAA GCACCCCGGC ACGGGAAAGG 960

CGGAGATTGT CCTGTTGGAC CATGGGCTTT ACCAGATGCT CACGGAAGAA TTCCGCCTGA 1020

ATTACTGCCA CCTCTGGCAG TCTCTGATCT GGACTGACAG GAAGAGAGTG AAGGAGTACA 1080

GCCAGCGACT GGGAGCCGGG GATCTCTACC CCTTGTTTGC CTGCATGCTG ACGGCGCGAT 1140

CGTGGGACTC GGTCAACAGA GGCATCAGCC AAGCTCCCGT CACTGCCACT GAGGACTTAG 1200

AGATTCGCAA CAACGCGGCC AACTACCTCC CCCAGATCAG CCATCTCCTC AACCACGTGC 1260

CGCGCCAGAT GCTGCTCATC TTGAAGACCA ACGACCTGCT GCGTGGCATT GAGGCCGCCC 1320

TGGGCACCCG CGCCAGCGCC AGCTCCTTTC TCAACATGTC ACGTTGCTGC ATCAGAGCGC 1380

TAGCTGAGCA CAAGAAGAAG AATACCTGTT CATTCTTCAG AAGGACCCAG ATCTCTTTCA 1440

GCGAGGCCTT CAACTTATGG CAGATCAACC TCCATGAGCT CATCCTGCGT GTGAAGGGGT 1500

TGAAGCTGGC TGACCGGGTC TTGGCCCTAA TATGCTGGCT GTTCCCTGCT CCACTCTGAG 1560

TGGAATTGCT CTCCCTGCCC CATTCTGGTG TCTTTCCACT CCTCAGCCCC TCATCTTGCC 1620

TCCACCCAGC TGCTCCATTT TTGCCACATC GTGGCCCGCA GCCCCAGAGT CACTGTCCAT 1680

GTCACCATCC TCCTCCTCCT TTGGAATCCT CTCCGCACAC TGTGGCCCTT GTCTCAGGGC 1740

CCACAAGCTG AACTGTGGCA TAGCTCTCTC TTCTTCTCCA AGAAGACTCA GCAGCCTACA 1800

TTCCCATTCC TGGTATGTGC CATTGGGTTG GATGTCCCCA CTACTTCCGT TAACCCTTCC 1860

CATTGTCAAG ATGTGCCACG GGTGCCACTG GGGGCACACT GAACTTGTAG GGAGTGTGAT 1920

TTTGTTGGAG GTGCACATGG TCTCTGAATT TGACAGAGAA CACCTTCCCT TTCCTTGCCA 1980

TGTCACCCTC CAGAGGAAGT CACACCTCAG CGAGGTGGTT TGGCATCTGG GGCCAACTCC 2040

ATTACAGCTA TGAGCTCACT GCTGTCAGTG ACGTTTGGTG TTTTCTGTAC TGTGTTTCAA 2100

TAAAAACTCC TTCAAGGTTG CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2160

AAAAAAAAAA A 2171

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Ala Arg Lys Ala Leu Lys Leu Ala Ser Trp Thr Ser Met Ala Leu
1 5 10 15

Ala Ala Ser Gly Ile Tyr Phe Tyr Ser Asn Lys Tyr Leu Asp Pro Asn
20 25 30

Asp Phe Gly Ala Val Arg Val Gly Arg Ala Val Ala Thr Thr Ala Val
35 40 45

```
Ile Ser Tyr Asp Tyr Leu Thr Ser Leu Lys Ser Val Pro Tyr Gly Ser
    50                  55                  60

Glu Glu Tyr Leu Gln Leu Arg Ser Lys Val His Leu Arg Ser Ala Arg
65                  70                  75                  80

Arg Leu Cys Glu Leu Cys Cys Ala Asn Arg Gly Thr Phe Ile Lys Val
                85                  90                  95

Gly Gln His Leu Gly Ala Leu Asp Tyr Leu Leu Pro Glu Glu Tyr Thr
            100                 105                 110

Ser Thr Leu Lys Val Leu His Ser Gln Ala Pro Gln Ser Ser Met Gln
        115                 120                 125

Glu Ile Arg Gln Val Ile Arg Glu Asp Leu Gly Lys Glu Val Leu Val
    130                 135                 140

Leu Ala Val Lys Gln Leu Phe Pro Glu Phe Glu Phe Met Trp Leu Val
145                 150                 155                 160

Asp Glu Ala Lys Lys Asn Leu Pro Leu Glu Leu Asp Phe Leu Asn Glu
                165                 170                 175

Gly Arg Asn Ala Glu Lys Val Ser Gln Met Leu Arg His Phe Asp Phe
            180                 185                 190

Leu Lys Val Pro Arg Ile His Trp Asp Leu Ser Thr Glu Arg Val Leu
        195                 200                 205

Leu Met Glu Phe Val Asp Gly Gly Gln Val Asn Asp Arg Asp Tyr Met
    210                 215                 220

Glu Arg Asn Lys Ile Asp Val Asn Glu Val Arg Ser Arg Ala Gln Gly
225                 230                 235                 240

Cys Cys Ala Gly Glu Arg Gly Leu Gly Gln Gly Cys Pro Gly Ser Ala
                245                 250                 255

Cys Val Ser Arg Ser His Ala Thr Trp Ala Arg Cys Ile Val Arg
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1613 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CATGGCGGCT CCCTTGGTCC TGGTGCTGGT GGTGGCTGTG ACAGTGCGGG CGGCCTTGTT 60

CCGCTCCAGT CTGGCCGAGT TCATTTCCGA GCGGGTGGAG GTGGTGTCCC CACTGAGCTC 120

TTGGAAGAGA GTGGTTGAAG GCCTTTCACT GTTGGACTTG GGAGTATCTC CGTATTCTGG 180

AGCAGTATTT CATGAAACTC CATTAATAAT ATACCTCTTT CATTTCCTAA TTGACTATGC 240

TGAATTGGTG TTTATGATAA CTGATGCACT CACTGCTATT GCCCTGTATT TTGCAATCCA 300

GGACTTCAAT AAAGTTGTGT TTAAAAAGCA GAAACTCCTC CTAGAACTGG AACAGTATGC 360

CCCAGATGTG GCCGAACTCA TCCGGACCCC TATGGAAATG CGTTACATCC CTTTGAAAGT 420

GGCCCTGTTC TATCTCTTAA ATCCTTACAC GATTTTGTCT TGTGTTGCCA AGTCTACCTG 480

TGCCATCAAC AACACCCTCA TTGCTTTCTT CATTTTGACT ACGATAAAAG TTTCATTATC 540

TGTAAAATGG GGACAGTAAT TGTACCCACT TCATGGAATT ATTGAGAAGA CTAAATGGCT 600

TAAGGCAGTG CTTTCCTCAG TGCTATTTTT CTTGCCTTAG CGACATACCA GTCTCTGAAC 660

CCACTCACCT TGTTTGTCCC AGGACTCCTC TATCTCCTCC AGCGGCAGTA CATACCTGTG 720

AAAATGAAGA GCAAAGCCTT CTGGATCTTT TCTTGGGAGT ATGCCATGAT GTATGTGGGA 780

AGCCTAGTGG TAATCATTTG CCTCTCCTTC TTCCTTCTCA GCTCTTGGGA TTTCATCCCC 840

GCAGTCTATG GCTTTATACT TTCTGTTCCA GATCTCACTC CAAACATTGG TCTTTTCTGG 900

TACTTCTTGG CAGAGATGTT TGAGCACTTC AGCCTCTTCT TTGTATGTGT GTTTCAGATC 960

AACGTCTTCT TCTACACCAT CCCCTTAGCC ATAAAGCTAA ATCCTGAGAA ACATCTTTGT 1020

CCTCACCTGC ATCATCATCG TCTGTTCCCT GCTCTTCCCT GTCCTGTGGC ACCTCTGGAT 1080

TTATGCAGGA AGTGCCAACT CTAATTTCTT TTATGCCATC ACACTGACCT TCAACGTTGG 1140

GCAGATCCTG CTCATCTCTG ATTACTTCTA TGCCTTCCTG CGGCGGGAGT ACTACCTCAC 1200

ACATGGCCTC TACTTGACCG CCAAGGATGG CACAGAGGCC ATGCTCGTGC TCAAGTAGGC 1260

CTGGCTGGCA CAGGGCTGCA TGGACCTCAG GGGGCTGTGG GGCCAGAAGY TGGGCCAAGC 1320

CCTCCAGCCA GAGTTGCCAG CAGGCGAGTG CTTGGGCAGA AGAGGTTCGA GTCCAGGGTC 1380

ACAAGTCTCT GGTACCAAAA GGGACCCATG GCTGACTGAC AGCAAGGCCT ATGGGGAAGA 1440

ACTGGGAGYT CCCCAACTTG GACCCCCACC TTGTGGCTCT GCACACCAAG GAGCCCCYTC 1500

CCAGACAGGA AGGAGAAGAG GCAGGTGAGC AGGGCTTGTT AGATTGTGGC TACTTAATAA 1560

ATGTTTTTTG TTATGAAGTC TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 1613

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 185 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Ala Ala Pro Leu Val Leu Val Leu Val Val Ala Val Thr Val Arg
1               5                  10                  15

Ala Ala Leu Phe Arg Ser Ser Leu Ala Glu Phe Ile Ser Glu Arg Val
            20                  25                  30

Glu Val Val Ser Pro Leu Ser Ser Trp Lys Arg Val Val Glu Gly Leu
        35                  40                  45

Ser Leu Leu Asp Leu Gly Val Ser Pro Tyr Ser Gly Ala Val Phe His
    50                  55                  60

Glu Thr Pro Leu Ile Ile Tyr Leu Phe His Phe Leu Ile Asp Tyr Ala
65                  70                  75                  80

Glu Leu Val Phe Met Ile Thr Asp Ala Leu Thr Ala Ile Ala Leu Tyr
                85                  90                  95

Phe Ala Ile Gln Asp Phe Asn Lys Val Val Phe Lys Lys Gln Lys Leu
            100                 105                 110

Leu Leu Glu Leu Glu Gln Tyr Ala Pro Asp Val Ala Glu Leu Ile Arg
        115                 120                 125

Thr Pro Met Glu Met Arg Tyr Ile Pro Leu Lys Val Ala Leu Phe Tyr
    130                 135                 140

Leu Leu Asn Pro Tyr Thr Ile Leu Ser Cys Val Ala Lys Ser Thr Cys
145                 150                 155                 160

Ala Ile Asn Asn Thr Leu Ile Ala Phe Phe Ile Leu Thr Thr Ile Lys
                165                 170                 175

Val Ser Leu Ser Val Lys Trp Gly Gln
            180                 185
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAACCCTGTC GGTCTTGGAG CGACGACGGC AGAACCAGGG TCCCTGGCGG TGCGGCGGGG 60

CCGGCGGGTG CAGCGGAAGC GGCGGCGGCG GCGGCAGTGA CGTCGCCGGG AACCCTAAGG 120

ACTCTGCAAT ATGAATAATT CCCTAGAGAA CACCATCTCC TTTGAAGAGT ACATCCGAGT 180

AAAGGCACGG TCTGTCCCGC AACACAGGAT GAAGGAATTT CTGGACTCAC TGGCCTCTAA 240

GGGGCCAGAA GCCCTTCAGG AGTTCCAGCA GACAGCCACC ACTACCATGG TGTACCAACA 300

GGGTGGGAAC TGCATATACA CAGACAGCAC TGAAGTGGCT GGGTCTTTGC TTGAACTTGC 360

CTGTCCAATC AC 372

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGGAAGCTC GAAATGGAGA AGGTGAACCT TATGACCCAG ATGTGCTCTA CTATATTTTC 60

CTGTGTATTC AAAAGTATCT TTTTGAAAAT GGAAGGGTAG ATGACATTTT CTCCGATCTT 120

TATTATGTTC GGTTCACGGA GTGGCTACAT GAAGTTCTGA AGGATGTTCA GCCCCGGGTC 180

ACTCCACTTG GCTATGTCTT GCCCAGCCAC GTGACTGAGG AGATGCTATG GGAGTGCAAG 240

CAGCTTGGGG CTCACTCCCC CTCCACCTTG CTGACCACCC TCATGTTCTT TAATACCAAG 300

TAAGTGTTCT AGAGGCTCCA CTGCTGGCAT CTGTCCAGTG AAGAGTGTGG AAACTATCCA 360

AGAGGCCTTC TGAATTCCTC TGACATATAT TTGAGAAACT GGGCTACTGA AAGCCCTAAC 420

CCCACTTGGC TGCATTTTAT TTGGTAACCA GTGAGGCAAA CACCCTTGCC AGACCCCTAC 480

CATCCATCTT GATGTGGTTC CTGCACTGGA CACTGCTTGG GTACGGGCCT GCCCAGATCT 540

TGGGAATGTG GGCAGTGGCT CCTCTGAAGC ACCAGTGGGC AGAGGATGAG TCATGGTATC 600

CT 602

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Trp Phe Leu His Trp Thr Leu Leu Gly Tyr Gly Pro Ala Gln Ile
1 5 10 15

Leu Gly Met Trp Ala Val Ala Pro Leu Lys His Gln Trp Ala Glu Asp
20 25 30

Glu Ser Trp Tyr Pro
35

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGGGAAAGGG CTTGGACTGT GAAAAGAAAT GTGGCCCCTT TCCATCTTCA AGAGAGATGG 60

AATTAATGAT GGATGGACCC TGGAGGGAAT CTCCCCAGCC GACTTCCACT GGGCTGACAG 120

ACTTTGCTGA CCACAGGGGA ACGATGTTCT TTTCTTTCTT CATGATCAGA CATAAACTTA 180

GCATTTTAAT GGAAGAAAAA TGAGGGGAAC TTCAATTATG ATTTATTAAA GACAATTTCT 240

ATTACACCCT CCTTTATGAC AAGTGACATT TTAGATGTAA AAGTAAAAAC TTTACCATGC 300

CTTTTTTTTT TTTGTTGGCC TAACATTGAG GCCTTAAAAC CTGAGGCTCC TGTGCCTGAT 360

GGAATTCTTG TAACATACAC TTGTGTATCA TATAAAGATA CCACTCTGTT TCTCTTATGT 420

ATTCTTACTC TAGTTGTTTA TTAAGAATGA CAAGCACGTC TTTTCAACAA AAAAAAAAAA 480

AAA 483

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1853 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CAGATTCGCT GCTGGAGTGC TGGATGGAGC CTTTCTCTGC CCTCTGTGAC ATTTCCAATT 60

TTAGATAATG CCTCACATCT CTGTCCCCCC GGGACCCCCT GGAGCCCCCA TGATCCCTAA 120

GAAGACAGCT TGAACCTAGA TCTCACCCCC AGGATGTTGC GGAGGCTGCT GGAGCGGCCT 180

TGCACGCTGG CCCTGCTTGT GGGCTCCCAG CTGGCTGTCA TGATGTACCT GTCACTGGGG 240

GGCTTCCGAA GTCTCAGTGC CCTATTTGGC CGAGATCAGG GACCGACATT TGACTATTCT 300

CACCCTCGTG ATGTCTACAG TAACCTCAGT CACCTGCCTG GGGCCCCAGG GGGTCCTCCA 360

GCTCCTCAAG GTCTGCCCTA CTGTCCAGAA CGATCTCCTC TCTTAGTGGG TCCTGTGTCG 420

GTGTCCTTTA GCCCAGTGCC ATCACTGGCA GAGATTGTGG AGCGGAATCC CCGGGTAGAA 480

CCAGGGGGCC GGTACCGCCC TGCAGGTTGT GAGCCCCGCT CCCGAACAGC CATCATTGTG 540

CCTCATCGTG CCCGGGAGCA CCACCTGCGC CTGCTGCTCT ACCACCTGCA CCCCTTCTTG 600

CAGCGCCAGC AGCTTGCTTA TGGCATCTAT GTCATCCACC AGGCTGGAAA TGGAACATTT 660

AACAGGGCAA AACTGTTGAA CGTTGGGGTG CGAGAGGCCC TGCGTGATGA AGAGTGGGAC 720

TGCCTGTTCT TGCACGATGT GGACCTCTTG CCAGAAAATG ACCACAATCT GTATGTGTGT 780

GACCCCCGGG GACCCCGCCA TGTTGCCGTT GCTATGAACA AGTTTGGATA CAGCCTCCCG 840

TACCCCCAGT ACTTCGGAAG AGTCTCAGCA CTTACTCCTG ACCAGTACCT GAAGATGAAT 900

GGCTTCCCCA ATGAATACTG GGGCTGGGGT GGTGAGGATG ACGACTTGCT ACCAGGGTGC 960

GCCTGGCTGG GATGAAGATC TCTCGGCCCC CCACATCTGT AGGACACTAT AAGATGGTGA 1020

AGCACCGAGG AGATAAGGGC AATGAGGAAA ATCCCCACAG ATTTGACCTC CTGGTCCGTA 1080

CCCAGAATTC CTGGACGCAA GATGGGATGA ACTCACTGAC ATACCAGTTG CTGGCTCGAG 1140

AGCTGGGGCC TCTTTATACC AACATCACAG CAGACATTGG GACTGACCCT CGGGGTCCTC 1200

GGGCTCCTTC TGGGCCACGT TACCCACCTG GTTCCTCCCA AGCCTTCCGT CAAGAGATGC 1260

TGCAACGCCG GCCCCCAGCC AGGCCTGGGC CTCTATCTAC TGCCAACCAC ACAGCCCTCC 1320

GAGGTTCACA CTGACTCCTC CTTCCTGTCT ACCTTAATCA TGAAACCGAA TTCATGGGGT 1380

TGTATTCTCC CCACCCTCAG CTCCTCACTG TTCTCAGAAG GATGTGAGGG AACTGAACTC 1440

TGGTGCCGTG CTAGGGGGTA GGGGCCTCTC CCTCACTGCT GGACTGGAGC TGGGCTCCTG 1500

TAGACCTGAG GGTCCNTCTY TCTAGGTCTC CTGTAGGGCT TAKGACTGTG AATCCTTGAT 1560

GTCATGATTT TATGTGACGA TTCCTAGGAG TCCCTGCCCC TAGAGTAGGA GCAGGGYTGG 1620

ACCCCAAGCC CNTCCYTYTT CCATGGAGAG AAGAGTGATC TGGYTTCTCC TCGGACCTCT 1680

GTGAATATTT ATTCTATTTA TGGTTCCCGG GAAGTTGTTT GGTGAAGGAA GCCCCTCCCC 1740

TGGGCATTTT CTGCCTATGC TGGAATAGCT CCCTCTTCTG GTCCTGGCTC AGGGGGCTGG 1800

GATTTTGATA TATTTTCTAA TAAAGGACTT TGTCTCGCAA AAAAAAAAAA AAA 1853

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 273 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Leu Arg Arg Leu Leu Glu Arg Pro Cys Thr Leu Ala Leu Leu Val
1 5 10 15

Gly Ser Gln Leu Ala Val Met Met Tyr Leu Ser Leu Gly Gly Phe Arg
20 25 30

Ser Leu Ser Ala Leu Phe Gly Arg Asp Gln Gly Pro Thr Phe Asp Tyr
35 40 45

Ser His Pro Arg Asp Val Tyr Ser Asn Leu Ser His Leu Pro Gly Ala

```
    50                  55                  60

Pro Gly Gly Pro Pro Ala Pro Gln Gly Leu Pro Tyr Cys Pro Glu Arg
65                  70                  75                  80

Ser Pro Leu Leu Val Gly Pro Val Ser Val Ser Phe Ser Pro Val Pro
                85                  90                  95

Ser Leu Ala Glu Ile Val Glu Arg Asn Pro Arg Val Glu Pro Gly Gly
            100                 105                 110

Arg Tyr Arg Pro Ala Gly Cys Glu Pro Arg Ser Arg Thr Ala Ile Ile
        115                 120                 125

Val Pro His Arg Ala Arg Glu His His Leu Arg Leu Leu Leu Tyr His
    130                 135                 140

Leu His Pro Phe Leu Gln Arg Gln Gln Leu Ala Tyr Gly Ile Tyr Val
145                 150                 155                 160

Ile His Gln Ala Gly Asn Gly Thr Phe Asn Arg Ala Lys Leu Leu Asn
                165                 170                 175

Val Gly Val Arg Glu Ala Leu Arg Asp Glu Glu Trp Asp Cys Leu Phe
            180                 185                 190

Leu His Asp Val Asp Leu Leu Pro Glu Asn Asp His Asn Leu Tyr Val
        195                 200                 205

Cys Asp Pro Arg Gly Pro Arg His Val Ala Val Ala Met Asn Lys Phe
    210                 215                 220

Gly Tyr Ser Leu Pro Tyr Pro Gln Tyr Phe Gly Arg Val Ser Ala Leu
225                 230                 235                 240

Thr Pro Asp Gln Tyr Leu Lys Met Asn Gly Phe Pro Asn Glu Tyr Trp
                245                 250                 255

Gly Trp Gly Gly Glu Asp Asp Asp Leu Leu Pro Gly Cys Ala Trp Leu
            260                 265                 270

Gly
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1686 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGATAAAGTA AGTGCTGTTT GGGCTAACAG GATCTCCTCT TGCAGTCTGC AGCCCAGGAC 60

GCTGATTCCA GCAGCGCCTT ACCGCGCAGC CCGAAGATTC ACTATGGTGA AAATCGCCTT 120

CAATACCCCT ACCGCCGTGC AAAAGGAGGA GGCGCGGCAA GACGTGGAGG CCCTCCTGAG 180

CCGCACGGTC AGAACTCAGA TACTGACCGG CAAGGAGCTC CGAGTTGCCA CCCAGGAAAA 240

AGAGGGCTCC TCTGGGAGAT GTATGCTTAC TCTCTTAGGC CTTTCATTCA TCTTGGCAGG 300

ACTTATTGTT GGTGGAGCCT GCATTTACAA GTACTTCATG CCCAAGAGCA CCATTTACCG 360

TGGAGAGATG TKCTTTTTTG ATTCTGAGGA TCCTGCAAAT TCCCTTCGTG GAGGAGAGCC 420

TAACTTCCTG CCTGTGACTG AGGAGGCTGA CATTCGTGAG GATGACAACA TTGCAATCAT 480

TGATGTGCCT GTCCCCAGTT TCTCTGATAG TGACCCTGCA GCAATTATTC ATGACTTTGA 540

AAAGGGAATG ACTGCTTACC TGGACTTGTT GCTGGGGAAC TGCTATCTGA TGCCCCTCAA 600

TACTTCTATT GTTATGCCTC CAAAAAATCT GGTAGAGYTC TTTGGCAAAC TGGCGAGTGG 660

CAGATATCTG CYTCAAACTT ATGTGGTTCG AGAAGACCTA GTTGCTGTGG AGGAAATTCG 720

TGATGTTAGT AACCTTGGCA TCTTTATTTA CCAACTTTGC AATAACAGAA AGTCCTTCCG 780

CCTTCGTCGC AGAGACCTCT TGCTGGGTTT CAACAAACGT GCCATTGATA AATGCTGGAA 840

GATTAGACAC TTCCCCAACG AATTTATTGT TGAGACCAAG ATCTGTCAAG AGTAAGAGGC 900

AACAGATAGA GTGTCCTTGG TAATAAGAAG TCAGAGATTT ACAATATGAC TTTAACATTA 960

AGGTTTATGG GATACTCAAG ATATTTACTC ATGCATTTAC TCTATTGCTT ATGCTTTAAA 1020

AAAAGGAAAA GAAAAAAACT ACTAACCACT GCAAGCTCTT GTCAAATTTT AGTTTAATTG 1080

GCATTGCTTG TTTTTTGAAA CTGAAATTAC ATGAGTTTCA TTTTTTCTTT GAATTTATAG 1140

GGTTTAGATT TCTGAAAGCA GCATGAATAT ATCACCTAAC ATCCTGACAA TAAATTCCAT 1200

CCGTTGTTTT TTTTGTTTGT TTGTTTTTTC TTTTCCTTTA AGTAAGCTCT TTATTCATCT 1260

TATGGTGCAG CAATTTTAAA ATTTGAAATA TTTTAAATTG TTTTTGAACT TTTTGTGTAA 1320

AATATATCAG ATCTCAACAT TGTTGGTTTC TTTTGTTTTT CATTTTGTAC AACTTTCTTG 1380

AATTTAGAAA TTACATCTTT GCAGTTCTGT TAGGTGCTCT GTAATTAACC TGACTTATAT 1440

GTGAACAATT TTCATGAGAC AGTCATTTTT AACTAATGCA GTGATTCTTT CTCACTACTA 1500

TCTGTATTGT GGAATGCACA AAATTGTGTA GGTGCTGAAT GCTGTAAGGA GTTTAGGTTG 1560

TATGAATTCT ACAACCCTAT AATAAATTTT ACTCTATAAA AAAAAAAAAA AAAAAAAAAA 1620

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1680

AAAAAA 1686

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Val Lys Ile Ala Phe Asn Thr Pro Thr Ala Val Gln Lys Glu Glu
1 5 10 15

Ala Arg Gln Asp Val Glu Ala Leu Leu Ser Arg Thr Val Arg Thr Gln
20 25 30

Ile Leu Thr Gly Lys Glu Leu Arg Val Ala Thr Gln Glu Lys Glu Gly
35 40 45

Ser Ser Gly Arg Cys Met Leu Thr Leu Leu Gly Leu Ser Phe Ile Leu
50 55 60

Ala Gly Leu Ile Val Gly Gly Ala Cys Ile Tyr Lys Tyr Phe Met Pro
65 70 75 80

Lys Ser Thr Ile Tyr Arg Gly Glu Met Xaa Phe Phe Asp Ser Glu Asp
85 90 95

Pro Ala Asn Ser Leu Arg Gly Gly Glu Pro Asn Phe Leu Pro Val Thr
100 105 110

Glu Glu Ala Asp Ile Arg Glu Asp Asp Asn Ile Ala Ile Ile Asp Val
115 120 125

Pro Val Pro Ser Phe Ser Asp Ser Asp Pro Ala Ala Ile Ile His Asp
130 135 140

Phe Glu Lys Gly Met Thr Ala Tyr Leu Asp Leu Leu Leu Gly Asn Cys
145 150 155 160

Tyr Leu Met Pro Leu Asn Thr Ser Ile Val Met Pro Pro Lys Asn Leu
165 170 175

Val Glu Xaa Phe Gly Lys Leu Ala Ser Gly Arg Tyr Leu Xaa Gln Thr

```
                180                 185                 190

Tyr Val Val Arg Glu Asp Leu Val Ala Val Glu Glu Ile Arg Asp Val
        195                 200                 205

Ser Asn Leu Gly Ile Phe Ile Tyr Gln Leu Cys Asn Asn Arg Lys Ser
    210                 215                 220

Phe Arg Leu Arg Arg Arg Asp Leu Leu Leu Gly Phe Asn Lys Arg Ala
225                 230                 235                 240

Ile Asp Lys Cys Trp Lys Ile Arg His Phe Pro Asn Glu Phe Ile Val
                245                 250                 255

Glu Thr Lys Ile Cys Gln Glu
            260
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TNCACATTCTC AGTGGGAACT TGATGAAC 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANATATAGGTG GAATGAATTC TATCCTTG 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GNTATAGTAAT AATAGCACAA AGGACGGG 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNGCCAGGAAA CCAAATCAAT GACTCTTT 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNTAATTGACG GTGTAGAGAA ATGAGGAA 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ANAAATGGAGC AGCTGGGTGG AGGCAAGA 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TNCGGAGATAC TCCCAAGTCC AACAGTGA 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GNTTAGGGCTT TCAGTAGCCC AGTTTCTC 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ANTGACAGGTA CATCATGACA GCCAGCTG 29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CNGGATGTTAG GTGATATATT CATGCTGC 29

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 264 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Gly Glu Val Lys Ser Ser Gly Leu Leu Cys Arg Gly Thr Trp Ala Trp
1               5                  10                  15

Ser Arg Leu Pro Arg Lys Cys Leu Cys Val Gln Ile Ser Arg His Leu
            20                  25                  30

Gly Lys Met Tyr Ser Glu Met Ile Phe Val Asn Gly Phe Val His Cys
        35                  40                  45
```

```
Asp Pro His Pro Gly Asn Val Leu Val Arg Lys His Pro Gly Thr Gly
    50                  55                  60

Lys Ala Glu Ile Val Leu Leu Asp His Gly Leu Tyr Gln Met Leu Thr
65                  70                  75                  80

Glu Glu Phe Arg Leu Asn Tyr Cys His Leu Trp Gln Ser Leu Ile Trp
                85                  90                  95

Thr Asp Arg Lys Arg Val Lys Glu Tyr Ser Gln Arg Leu Gly Ala Gly
            100                 105                 110

Asp Leu Tyr Pro Leu Phe Ala Cys Met Leu Thr Ala Arg Ser Trp Asp
        115                 120                 125

Ser Val Asn Arg Gly Ile Ser Gln Ala Pro Val Thr Ala Thr Glu Asp
    130                 135                 140

Leu Glu Ile Arg Asn Asn Ala Ala Asn Tyr Leu Pro Gln Ile Ser His
145                 150                 155                 160

Leu Leu Asn His Val Pro Arg Gln Met Leu Leu Ile Leu Lys Thr Asn
                165                 170                 175

Asp Leu Leu Arg Gly Ile Glu Ala Ala Leu Gly Thr Arg Ala Ser Ala
            180                 185                 190

Ser Ser Phe Leu Asn Met Ser Arg Cys Cys Ile Arg Ala Leu Ala Glu
        195                 200                 205

His Lys Lys Lys Asn Thr Cys Ser Phe Phe Arg Arg Thr Gln Ile Ser
    210                 215                 220

Phe Ser Glu Ala Phe Asn Leu Trp Gln Ile Asn Leu His Glu Leu Ile
225                 230                 235                 240

Leu Arg Val Lys Gly Leu Lys Leu Ala Asp Arg Val Leu Ala Leu Ile
                245                 250                 255

Cys Trp Leu Phe Pro Ala Pro Leu
            260
```

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 54 to nucleotide 737;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 188 to nucleotide 671;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bf171_6 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bf171_6 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 109 to amino acid 118 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 46 to amino acid 206;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 109 to amino acid 118 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone bf171_6 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 46 to amino acid 206.

11. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 135 to nucleotide 1169;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 875;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ck181_7 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 167 to amino acid 176 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

15. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 247;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 167 to amino acid 176 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone ck181_7 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 882 to nucleotide 1106;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1050 to nucleotide 1106;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1028 to nucleotide 1395;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone co736_3 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone co736_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 32 to amino acid 41 of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

18. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 32 to amino acid 41 of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone co736_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2283 to nucleotide 2858;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1164 to nucleotide 1433;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dm26_2 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 91 to amino acid 100 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

21. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 91 to amino acid 100 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone dm26_2 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 168 to nucleotide 683;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 318 to nucleotide 683;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eq229_3 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 81 to amino acid 90 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 53 to amino acid 172;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 81 to amino acid 90 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone eq229_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9 and SEQ ID NO:11.

26. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 67 to nucleotide 879;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 118 to nucleotide 879;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1224 to nucleotide 2171;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh3_6 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

27. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 119;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh3_6 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:12.

29. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 2 to nucleotide 556;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 53 to nucleotide 556;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 367;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fs87_3 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 87 to amino acid 96 of SEQ ID NO:15;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 87 to amino acid 96 of SEQ ID NO:15; and

(c) the amino acid sequence encoded by the cDNA insert of clone fs87_3 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:14.

32. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 492 to nucleotide 602;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

(e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fy530_2 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:18;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

33. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:18; and

(c) the amino acid sequence encoded by the cDNA insert of clone fy530_2 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17, SEQ ID NO:16, and SEQ ID NO:19 .

35. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 154 to nucleotide 972;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 341;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ge51_1 deposited under accession number ATCC 98371;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:21;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

36. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 62;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone ge51_1 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

38. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 104 to nucleotide 892;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 299 to nucleotide 892;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 798 to nucleotide 1261;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gx183_1 deposited under accession number ATCC 98371;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity, the fragment comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:23;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

39. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 53 to amino acid 89;

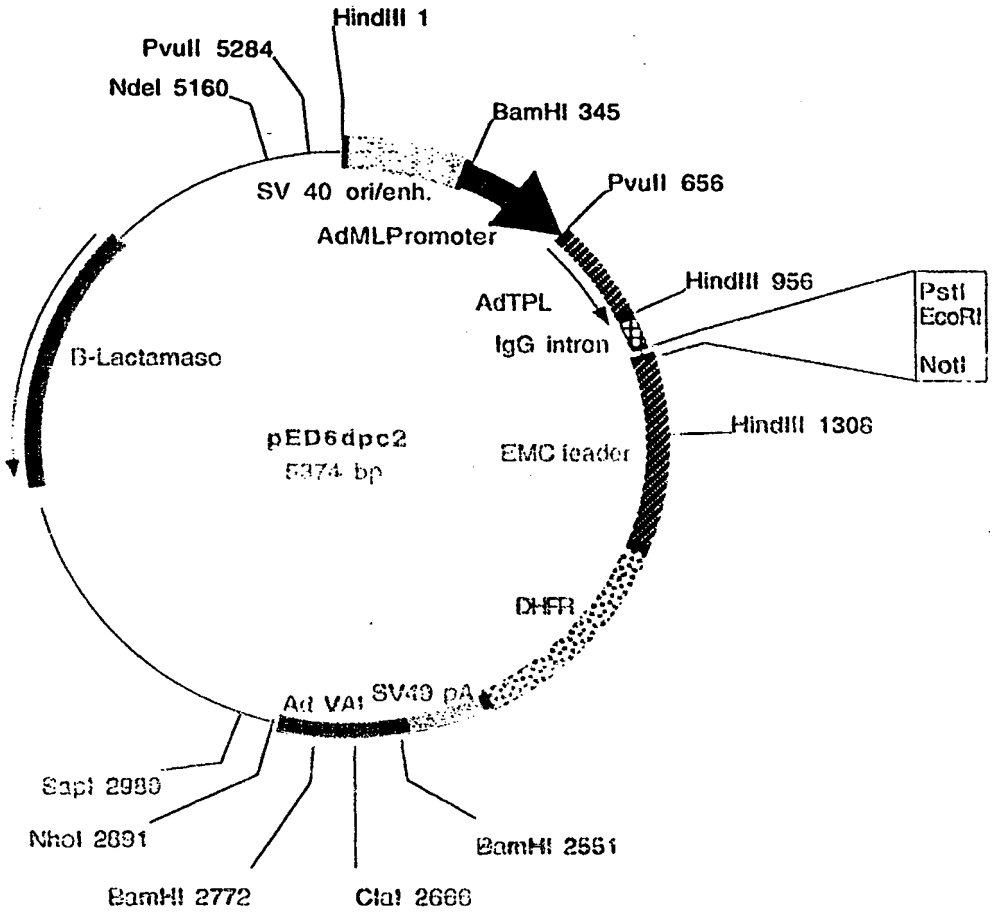
(c) fragments of the amino acid sequence of SEQ ID NO:23 comprising the amino acid sequence from amino acid 126 to amino acid 135 of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone gx183_1 deposited under accession number ATCC 98371;

the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:22.

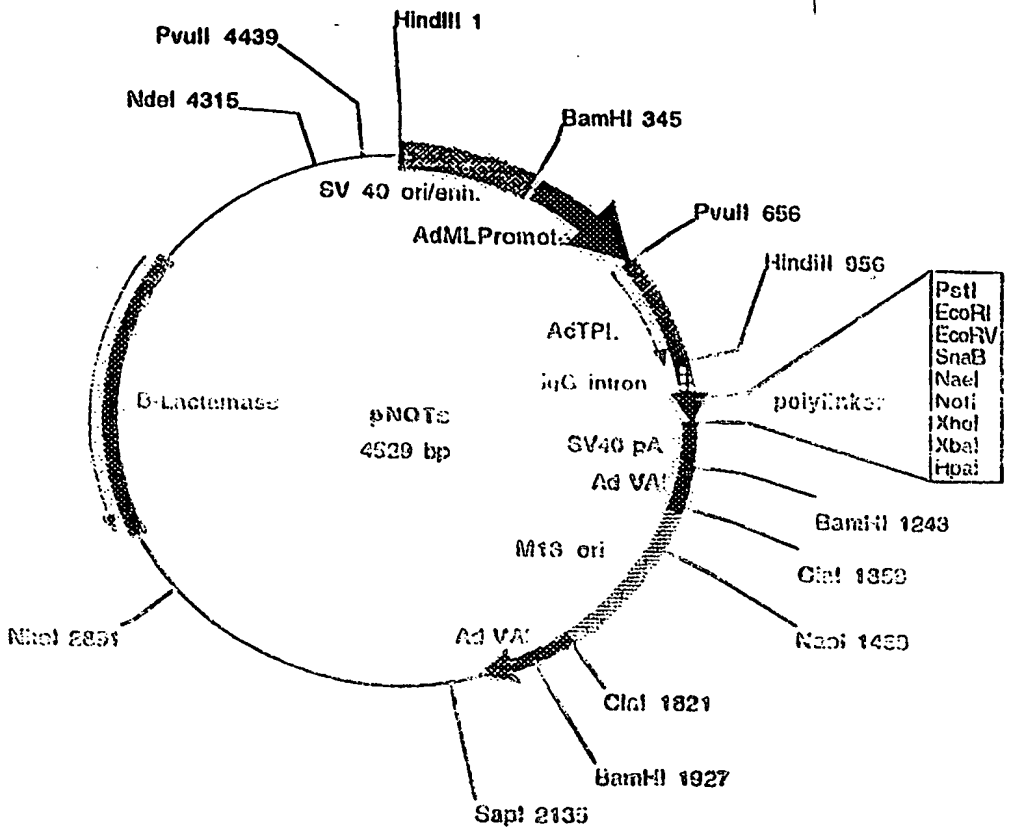
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, A61K 38/17

A3

(11) International Publication Number: WO 98/42741

(43) International Publication Date: 1 October 1998 (01.10.98)

(21) International Application Number: PCT/US98/05972

(22) International Filing Date: 25 March 1998 (25.03.98)

(30) Priority Data:

08/825,145	25 March 1997 (25.03.97)	US
09/046,881	24 March 1998 (24.03.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Street, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
23 December 1998 (23.12.98)

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/05972

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	FILIPEK A. AND WOJDA U.: "p30, a novel protein target of mouse calcyclin (S100A6)" BIOCHEMICAL JOURNAL, vol. 320, no. 2, 1 December 1996, pages 585-587, XP002069354 see page 587, left-hand column, paragraph 1 ---	1-13
X	Database EMBL, entry HSAA47377, Accession number AA147377 14 December 1996 99% identity with Seq.ID:1 nt.5-348 XP002069355 cited in the application ---	1,13
Y	see the whole document ---	1-13
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 June 1998	28.09.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/05972

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL, entry HSAA29803, Accession number AA129803 6 December 1996 99% identity with Seq.ID:1 nt.202-603 XP002069356	1,13
Y	see the whole document	1-13

X	Database EMBL, entry HSAA47341, Accession number AA147341 14 December 1996 97% identity with Seq.ID:1 nt.509-937 reverse orientation XP002069357 see the whole document	1,13

X	Database EMBL, entry HSAA28382, Accession number AA128382 6 December 1996 99% identity with Seq.ID:1 nt.614-1045 reverse orientation XP002069358 see the whole document	1,13

X	Database EMBL, entry HSAA59926, Accession number AA159926 19 January 1997 99% identity with Seq.ID:1 nt.973-1459 reverse orientation XP002069359 see the whole document	1,13

X	EP 0 679 716 A (MATSUBARA KENICHI; OKUBO KOUSAKU (JP)) 2 November 1995 see abstract see page 5, line 34 - page 9, line 45 see page 195; table 175 Seq.ID:6271 is 100% homologous to Seq.ID1 nt.609-891 see page 1783 - page 1784	1,13

A	WO 97 07198 A (GENETICS INSTITUTE INC.) 27 February 1997	

A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 cited in the application	

INTERNATIONAL SEARCH REPORT

Intr tional application No.
PCT/US 98/05972

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-13

   Polynucleotide comprising the nucleotide sequence of Seq.ID:1 and encoding a polypeptide of Seq.ID:2 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Host cell transformed with said polynucleotide. Protein comprising an amino acid sequence of Seq.ID:2 or fragments thereof. Process for producing said protein. Application of said protein in therapy.

2. Claims: 14-16

   Polynucleotide comprising the nucleotide sequence of Seq.ID:3 and encoding a polypeptide of Seq.ID:4 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Protein comprising an amino acid sequence of Seq.ID:4 or fragments thereof.

3. Claims: 17-19

   As invention 2 but concerning Seq.ID:5 and 6.

4. Claims: 20-22

   As invention 2 but concerning Seq.ID:7 and 8.

5. Claims: 23-25

   As invention 2 but concerning Seq.ID:9, 10 and 11.

6. Claims: 26-28

   As invention 2 but concerning Seq.ID:12 and 13.

7. Claims: 29-31

   As invention 2 but concerning Seq.ID:14 and 15.

8. Claims: 32-34

   As invention 2 but concerning Seq.ID:16, 17, 18 and 19.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

9. Claims: 35-37

As invention 2 but concerning Seq.ID:20 and 21.

10. Claims: 38-40

As invention 2 but concerning Seq.ID:22 and 23.

INTERNATIONAL SEARCH REPORT

In ...iation on patent family members

International Application No
PCT/US 98/05972

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0679716 A	02-11-95	AU 8116494 A	13-06-95
		CA 2153480 A	01-06-95
		WO 9514772 A	01-06-95
WO 9707198 A	27-02-97	US 5707829 A	13-01-98
		AU 6712396 A	18-02-97
		AU 6768596 A	12-03-97
		EP 0839196 A	06-05-98
		EP 0851875 A	08-07-98
		WO 9704097 A	06-02-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98